

IN THE SPECIFICATION

Enclosed herewith is a Substitute Specification.



#8
L5B
9/29/04

RECEIVED

AUG 30 2004

Technology Center 2600

SUBSTITUTE SPECIFICATION

TITLE: SYSTEM FOR CONTROLLING AN APPARATUS WITH SPEECH
COMMANDS

INVENTOR(S): CHIH-CHUAN YEN ET AL.

SERIAL NO.: 09/912,133

ATTY. DOCKET NO.: PHTW 000005

SYSTEM FOR CONTROLLING AN APPARATUS WITH SPEECH COMMANDS

BACKGROUND OF THE INVENTION

Field Of The Invention

[0001] The invention relates to a system including an apparatus and a remote control for controlling the apparatus, the system
5 further including a speech processor for processing speech commands, and the remote control comprising a microphone for enabling a user of the remote control to input the speech commands. The invention further relates to a remote control and an apparatus for use in the above system. The invention further relates to a
10 method of controlling an apparatus, comprising a step of processing speech commands for controlling said apparatus.

Description Of The Related Art

[0002] Recent developments in speech recognition techniques have
15 enabled users of electronic systems to control these systems by spoken commands. Often, such systems are used in a multi-user setting, e.g., a TV system is used by different members of a household. One of the problems with such systems is that two or more users may address the system at the same time. Since the users
20 can utter the commands at any time and location, it is very difficult to resolve the input conflict. Even if the speakers utter their commands at different times, contradictory commands from different users still cause a problem. Obviously, discrimination of

speech input is a serious concern for voice control with multiple users.

[0003] U.S. Patent 5,777,571 discloses a solution to this problem. Users are registered as authorized users, and their speech
5 commands are distinguished from speech commands from other users by means of voice identification techniques. A problem with this solution is that voice identification is an expensive and still unreliable technique, and new or occasional users need to be introduced to the system for authorization. Such an introduction
10 often involves a lengthy training phase. Furthermore, this solution cannot resolve the problem that many speakers speak at the same time.

SUMMARY OF THE INVENTION

15 [0004] It is an object of the invention to provide an improved system and method of the type defined in the opening paragraph. To that end, the system also comprises a further microphone for enabling further users of the system to input speech commands. The system according to the invention thus provides (at least) two
20 microphones for controlling the apparatus. One of the microphones is located on the remote control and is arranged to pick up speech commands uttered by the user of the remote control. The other microphone is located elsewhere, e.g., on the apparatus or at a central place in the room, and is arranged to pick up speech
25 commands uttered by other users of the system who are not currently

operating the remote control. In this way, it is achieved that the system can distinguish speech commands from the user operating the remote control, on the one hand, and other users, on the other hand. This guarantees that the speech commands uttered by the user who is operating the remote control are optimally recognized, since the microphone is located relatively close to the user and can have suitable characteristics to pick-up sounds from the appropriate direction only. When multiple users are speaking simultaneously, the system may give priority to signals received from the microphone on the remote control, so that at least commands uttered by the user of the remote control will be recognized and processed.

[0005] An embodiment of the system according to the invention is characterized in that the further microphone is an omnidirectional microphone. In this way, it is achieved that the positions of the other users in the room are not critical. By contrast, the microphone on the remote control is preferably unidirectional and, when held in a normal manner, oriented so as to be aimed at the user's mouth.

[0006] An embodiment of the system according to the invention is characterized in that the system comprises input designation means for enabling the user to selectively designate the microphone and/or the further microphone as a signal source to the speech processor. The system thus enables the speech commands obtained from the microphone and the further microphone to be selectively transmitted to the speech processor. For example, the speech

processor may be controlled to process speech input from the microphone on the remote control only. Alternatively, the speech processor may be controlled to process speech input from the further microphone only, or from both the further microphone and the microphone on the remote control. Finally, the speech processor may be decoupled from both microphones, thus disabling the speech command facility. Preferably, the input designation means can be controlled via the remote control. For example, the remote control may comprise a button enabling the user to switch between three different designations, e.g., accept no speech processing at all, accept speech commands from the microphone on the remote control only, and accept speech commands from the further microphone only. Alternatively, speech commands obtained from the further microphone may be accepted by default, while pressing a button on the remote control causes the microphone on the remote control to be temporarily designated as the input source of the speech processor.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] These and other aspects of the invention are apparent from and will be elucidated, by way of a non-limitative example, with reference to the embodiment(s) described hereinafter. In the drawings:

[0008] Figure 1 shows a television receiver and a remote control in an embodiment of a system according to the invention; and

[0009] Figure 2 shows a block schematic diagram of a television receiver in an embodiment of an apparatus according to the invention.

5

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0010] Figure 1 shows a television receiver 101 and a remote control 102 as an embodiment of a system according to the invention. The remote control 102 comprises an infrared (IR) transmitter 103, a microphone 104 and control elements including, inter alia, a button 105. The television receiver 101 comprises an IR receiver 106 and a further microphone 107. The television receiver can be operated by the remote control 102. For that purpose, the remote control 102 has a plurality of keys for controlling various functions of the television receiver 101. For example, the remote control 102 may have numerical keys, zap keys, etc., which are well known in the art. Additionally, a user of the remote control 102 can enter speech commands via the microphone 104, which are then transmitted to the IR receiver 106 of the television receiver 101 and converted to corresponding control commands by a speech processor, described hereinafter. The further microphone 107 is an omnidirectional microphone which picks up speech signals from any direction, thus enabling other users, who are not currently holding the remote control 102, to control the television receiver 101 by voice commands. The button 105 controls input designation means, described hereinafter, for designating the

microphone 104 and/or the further microphone 107 as a signal source for the speech processor.

[0011] Figure 2 shows diagram of a television receiver as an embodiment of an apparatus according to the invention. For consistency and ease of understanding, the same reference numerals as in Figure 1 are used for items having functions similar to those presented in Figure 1. TV signals are received over the air by an antenna 201 or, alternatively, from a cable network. One of the TV signals is selected by a tuner 202, decoded and split into an audio signal, a video signal and a data signal. The audio signal is further processed by an audio processor 203 and a loudspeaker 204. The video signal is further processed by a video processor 205 and presented on a screen 206. The data signal is supplied to a central processing unit (hereinafter "CPU") 208, which comprises one or more microprocessors capable of executing program instructions. These program instructions comprise parts of software modules including, inter alia, a command interpreter 209, and a speech processor 211. The CPU 208 is capable of controlling functions of the television receiver, as well as transmitting data to the video processor 205 to be presented on the screen 206. The command interpreter 209 receives user commands from the IR receiver 106, which, in turn, receives IR signals from the remote control 102, and supplies them to the CPU 208 to be processed. For example, when the user enters a channel number, the CPU 208 controls the tuner 202 to select the corresponding channel, and sends data to the

video processor 205 to be displayed on the screen 206, e.g., the preset number, the channel name and the program category are displayed for a few seconds. When the user issues a zap-command, e.g., by pressing either of the up/down keys on the remote control, the similar data is displayed, and the tuner 202 is controlled to select a channel which follows or precedes the current channel.

[0012] The CPU 208 is further capable of controlling a switch 215 which constitutes input designation means for the speech processor 211. The switch 215 can adopt three states. A first state of the switch 215 designates the microphone 104 of the remote control 102 as the signal source for the speech processor 211. Speech signals from the microphone 104 are converted into IR signals, received by the IR receiver 106 and supplied to the switch 215. In the first state of the switch 215, the speech signals thus obtained from the microphone 104 are input to the speech processor 211 and converted into control commands which are then supplied to the command interpreter 209.

[0013] A second state of switch 215 designates the microphone 107 as the signal source for the speech processor 211. The microphone 107 is a part of the television receiver 101, and signals obtained from the microphone 107 are supplied directly to a second contact of the switch 215.

[0014] A third state of switch 215 designates neither the microphone 104 nor the microphone 107 as the signal source for the speech processor 211, thus disabling speech input completely.

[0015] The switch 215 is controlled by the CPU 208 in response to control signals received from the IR receiver 106, these control signals being, in turn, generated in response to the user of the remote control 102 operating the button 105. After system
5 initialization, the switch 215 assumes the first state as described above and depicted in Figure 2. If the user presses the button 105, the switch 215 assumes the second state, thus designating the microphone 107 as the signal source to the speech processor 211 and disabling the microphone 104. If the user presses the button 105 a
10 second time, the switch 215 assumes the third state, thus designating neither the microphone 104 nor the microphone 107. Pressing the button 105 again returns the switch 215 to the first state.

[0016] In an alternative embodiment, the switch 215 can assume a
15 fourth state, wherein signals from both microphones are accepted. However, if signals are received from both microphones simultaneously, signals received from the microphone 107 are disregarded in favor of the signals received from the microphone 104.

20 [0017] In summary, the invention relates to a system comprising a speech processor for controlling an apparatus (101) with speech commands. The system, according to the invention, includes a remote control (102) having a microphone (104) for the input speech commands. The system also includes a further microphone (107) for
25 enabling other users of the system to also issue speech commands.

The system may have input designation means (105) for enabling the user to selectively designate the microphone (104) and/or the further microphone (107) as a signal source to the speech processor.

- 5 [0018] Although the invention has been described with reference to particular illustrative embodiments, variants and modifications are possible within the scope of the inventive concept. Thus, for example, the speech commands may be transmitted from the remote control by radio frequency (RF) signals instead of IR signals.
- 10 Furthermore, instead of, or in addition to, being included in the controlled apparatus, a speech processor may be included in the remote control. The input designation means may be controlled by a control element on the remote control or on the controlled
- 15 apparatus. The control element may be a single-state toggle button as described above, or any other appropriate control element, such as a 'radio button' for each state, or a multi-position switch, each position of which corresponding to a particular state. The input designation means could be a switch as described
- 20 hereinbefore, or a more sophisticated switching circuit under control of the speech processor itself. This would enable the input designation means to be controlled by dedicated speech commands irrespective of the currently selected state. The speech processor would then first try to detect such dedicated speech commands
- 25 irrespective of whether it is allowed to accept speech commands from the respective microphone, and subsequently adjust the input

designation in accordance with the dedicated speech commands.

Preferably, only a user of the remote control would be allowed to control the input designation means in such manner.

[0019] More than two microphones may be used, which may be
5 located at various positions within the system, e.g., on the
apparatus and the remote control as already described, or in the
vicinity of the controlled apparatus. The system may also comprise
multiple remote controls, each comprising a microphone. Preferably,
one of the remote controls serves as a master control and is the
10 only remote control capable of controlling the input designation
means.

[0020] The use of the verb 'comprise' does not exclude the
presence of any elements or steps other than those defined in a
claim. In the claims, any reference signs placed between
15 parentheses shall not be construed as limiting the claim. The
invention can be implemented by means of hardware comprising
several distinct elements, and by means of a suitably programmed
computer. In claims in which several means are defined, several of
these means can be embodied by one and the same item of hardware.

20